



# Learning to Play Bayesian Games

## Citation

Dekel, Eddie, Drew Fudenberg, and David K. Levine. 2004. Learning to play Bayesian games. Games and Economic Behavior 46, no. 2: 282-303.

## Published Version

[http://dx.doi.org/10.1016/S0899-8256\(03\)00121-0](http://dx.doi.org/10.1016/S0899-8256(03)00121-0)

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:3200612>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

# Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# **Learning to Play Bayesian Games<sup>1</sup>**

Eddie Dekel, Drew Fudenberg and David K. Levine<sup>2</sup>

First draft: December 23, 1996

Current revision: July 22, 2002

## **Abstract**

This paper discusses the implications of learning theory for the analysis of games with a move by Nature. One goal is to illuminate the issues that arise when modeling situations where players are learning about the distribution of Nature's move as well as learning about the opponents' strategies. A second goal is to argue that quite restrictive assumptions are necessary to justify the concept of Nash equilibrium without a common prior as a steady state of a learning process

---

<sup>1</sup> This work is supported by the National Science Foundation under Grants 99-86170, 97-30181, and 97-30493. We are grateful to Pierpaolo Battigalli, Dan Hojman and Adam Szeidl, a referee, and an Associate Editor for helpful comments.

<sup>2</sup> Departments of Economics: Northwestern and Tel Aviv Universities, Harvard University and UCLA.

## 1. Introduction

This paper discusses the implications of learning theory for the analysis of games with a move by Nature. Our premise is that equilibrium in games arises as the result of learning, and that just what people will learn depends both on the true distribution of Nature's move and on what they observe when the game is played. One of our goals is to illuminate some of the issues involved in modeling players' learning about opponents' strategies when the distribution of Nature's moves is also unknown. In this vein, we show how the relevant equilibrium concept changes when there are many agents in the role of a single "player," and when Nature's move is determined once and for all at the beginning of the game instead of being drawn independently each period. A second and more specific goal is to investigate the concept of Nash equilibrium without a common prior, in which players have correct and hence common beliefs about one another's strategies, but disagree about the distribution over Nature's moves.<sup>3</sup> This solution is worth considering given the recent popularity of papers that apply it, such as Banerjee and Somanathan [2001], Piketty [1995], and Spector [2000].<sup>4</sup> We argue that Nash equilibrium without a common prior is difficult to justify as the long-run result of a learning process, because it takes very special assumptions for the set of such equilibria to coincide with the set of steady states that could arise from learning.

The intuition for our concern is simple: In order for repeated observations to lead players to learn the distribution of opponents' strategies, the signals observed at the end of each round of play must be sufficiently informative. Such information will tend to

---

<sup>3</sup> Harsanyi [1967-8] proposed that games of incomplete information be analyzed as games with a move by Nature, where the move by Nature is an artificial construction (see, e.g., Dekel and Gul [1997] for a discussion). He defined (Bayesian) Nash equilibrium in his model both for the case of a common prior and for the case where priors are allowed to differ; this is the notion of Nash equilibrium we use herein. Our approach is motivated by a learning model in which play is repeated. In those cases where we assume that the move by Nature is also repeatedly drawn and its consequences observed, the interpretation of Nature's move as an artificial construction is not necessary, so it seems inappropriate there.

<sup>4</sup> We do not explore these applications in detail, so in particular we do not claim that their use of Nash equilibrium is inappropriate. We only want to argue that in the context of incorrect priors, the use of Nash equilibrium requires more careful justification than is typically given. In fact, Spector [2000] assumes that actions are observed while payoffs are not, noting that, while these are fairly extreme assumptions, if payoffs were observed then the players would learn the true distribution of Nature's move.

lead players to also have correct and hence identical beliefs about the distribution of Nature's moves. While this basic argument is straightforward, our examples highlight some less obvious points.

Our point of departure is the notion of self-confirming equilibrium, which has been used to model the outcome of learning processes (see, e.g., references in footnotes 5, 7, and 8). This equilibrium concept requires that each player's strategy is optimal given the player's beliefs about opponents' strategies and Nature's moves, and that these beliefs are consistent with the distribution of outcomes in the game that the player observes. In Section 2 we give a formal definition that extends the notion to allow for different possible maps from outcomes of the game to observations, in particular allowing for the case where players do not observe the actions taken by others or their own payoffs.

When players observe the "outcome" of round of play, meaning both the actions taken and the realization of Nature's move, the set of self-confirming equilibria is the same as the set of Nash equilibria with a common prior, so any strategy profile that is not a Nash equilibrium with a common prior cannot be a self-confirming equilibrium. When players observe less than the outcome, more beliefs are consistent with the players' observations, so a Nash equilibrium without a common prior may be a self-confirming equilibrium, but only because the set of self-confirming equilibria is then large (and therefore may well include many outcomes that are not Nash equilibria). This leads us, in Section 3, to explore the relationship between the set of Nash equilibria when players have given, possibly inconsistent, beliefs about Nature on the one hand, and the set of self-confirming equilibria with the same beliefs on the other; we use a series of examples and simple propositions to illustrate our concern with the use of Nash equilibrium when the prior is not common. Propositions 1, 2, and 3 present cases in which these sets only coincide for the case of a common prior. Proposition 4 presents a case where there are self-confirming equilibria with correct beliefs about Nature that are not Nash even with *correct* prior. On the other hand, Proposition 5 identifies a very special case where the sets of Nash and self-confirming equilibria do coincide.

Section 4 shows how to extend the equilibrium concept used in Section 3 to games in which there are many agents in the role of each player, and agents are randomly

matched each period. This situation is of interest because it serves to motivate our assumption that subjects ignore repeated-game considerations and try to maximize their payoff in each play of the stage game; for the same reason, this is a commonly used design in game theory experiments. Section 5 considers alternate specifications of the stochastic structure of Nature's move, explaining how they alter the definition of self-confirming equilibrium and its properties. We begin with the case where Nature's to make a once-and-for-all choice of a single profile of types that will apply to all matches; this specification has been used in many game-theory experiments. We then consider the case where Nature makes a once-and-for-all choice for each agent, but where different agents in a given player role can have different types, and we conclude with a brief discussion of alternate, more complex, stochastic structures.

Of course, our conclusions are all based on the idea of learning as a justification for equilibrium. In particular, if players achieve equilibrium through deliberation, then Nash equilibrium without common priors may be sensible. However, it is not at all clear how this would work in the presence of multiple equilibria. Moreover, the practical difficulties in computation and the amount of information that can be required to find an equilibrium makes this seem to us a weak justification for studying Nash equilibrium.

Our criticism of Nash equilibrium without a common prior does not mean we are arguing for the common prior assumption *per se*, and indeed we are sympathetic to the idea that some situations are better modeled without common priors; our concerns are with the uncritical use of the Nash equilibrium solution concept in the absence of a common prior. Our learning-theoretic critique is related to two other problems of Nash equilibrium without a common prior. One is internal consistency: a Nash equilibrium when players have different priors in general is not a Nash equilibrium when Nature is replaced with a player who is indifferent among all her choices and who behaves exactly as did Nature, because in a Nash equilibrium the strategy of the player replacing Nature is known. A related problem (Dekel and Gul [1997]) is that the epistemic foundations of Nash equilibrium without a common prior are unappealing. The epistemic foundation for Nash equilibrium relies on a common prior about strategies, and it is not obvious why we

should impose this on the states of Nature underlying the strategic uncertainty and not on those corresponding to other aspects of the state of Nature.

As mentioned, this paper is related to other work that uses notions similar to self-confirming equilibrium. Jordan [1993] and Jackson and Kalai [1997] are also related. Like us, they compare the long-run outcomes of a learning process to the equilibria of a static stage game, and several of our findings are similar to some of theirs. One important difference between those papers and this one is that they study Bayesian equilibrium play of an overall “recurring game” that has a common prior, while we study the steady states of non-equilibrium learning processes that need not start out with common beliefs. We say more about the relationship between these papers in sections 3 and 5.

## 2. The Model

We consider a static simultaneous-move game with  $I$  player roles. (All parameters of the game, including the number of players, and their possible actions and types, are assumed to be finite.) In the static game, Nature moves first, determining players’ types, which we denote  $\theta_i \in \Theta_i$ . To model cases where the types alone do not determine the realized payoffs, we also allow Nature to pick  $\theta_0 \in \Theta_0$ ; we call this “Nature’s type.” Players observe their types, and then simultaneously choose actions  $a_i \in A_i$  as a function of their type, so that a strategy  $\sigma_i$  for player  $i$  is a map from her types to mixed actions. Player  $i$ ’s utility  $u_i(a, \theta)$  depends on the profile  $a = (a_1, \dots, a_I) \in A$  of realized actions, and on the realization  $\theta = (\theta_0, \theta_1, \dots, \theta_I) \in \Theta$  of Nature’s move. When player  $i$ ’s utility does not depend on Nature’s move beyond  $i$ ’s own type, *i.e.*,  $u_i(a, \theta) = u_i(a, \theta_i)$ , we refer to the game as having *private values*. For any finite set  $X$ , we let  $\Delta(X)$  denote the space of probability distributions over  $X$ . Player  $i$ ’s (stage-game) prior about Nature’s move is denoted  $\mu^i \in \Delta(\Theta)$ ;  $\mu = \{\mu^1, \dots, \mu^I\}$  is the profile of (stage-game) priors. When  $\mu^i = \mu^j$  for all  $i$  and  $j$ , the game has a *common prior*; in the complementary case where  $\mu^i \neq \mu^j$  for some  $i$  and  $j$  we say that the priors are *diverse*.

Our solution concept is motivated by thinking about a learning environment in which the game given above is played repeatedly. We suppose that players know their own payoff functions and the sets of possible moves of all of the players ( $A$ ). We suppose that the set  $\Theta$  contains all of the type profiles that any player thinks is possible; some of

these may have probability 0 under the true distribution governing Nature's move. Players are not assumed to know either the strategies used by other players or the distribution of Nature's move; they learn about these variables from their observations after each period of play. For the time being, we also suppose that each period the types are drawn independently over time from a fixed distribution  $p$ ; this is the simplest specification, and as we will see, it is the one that corresponds most closely to standard models of Bayesian games. The distribution  $p$  corresponds to the true distribution of Nature's move in the stage game, so when  $\mu^i = p$  for all players  $i$  we say that *the priors are correct*.<sup>5</sup> For the time being, we also assume that there is a single agent in each player role. Sections 4 and 5 discuss the case where there is a large population of agents in each role who are matched together to play the game.

Of course, what players might learn from repeated play depends on what they observe at the end of each round of play. To model this, we adopt the formalism used by Battigalli [1987] and Rubinstein and Wolinsky [1994], and suppose that after each play of the game, players receive private signals  $y_i = y_i(a, \theta)$  which is their only information about Nature's and their opponents' moves. It is natural to assume that players observe their own actions and types, but whether or not they observe others' actions, or their own and others' payoffs, depends on the observation structure and will affect which outcomes can arise in a steady state. We assume that each player observes her own private signal  $y_i$ , along with her own action and own type.<sup>6</sup> Since the definition given above places no restrictions on the signal functions, it is consistent with the signal being the terminal node of an extensive-form game, so can be applied to more than just static games with an initial move by Nature. However, to focus on the issue of games with moves by Nature, we will not consider an explicit extensive-form structure, and we concentrate on

---

<sup>5</sup> Note that if players are Bayesians they will have a prior about the state of the overall learning process, and this prior need not be the fixed  $\mu^i$  that is taken as data in the specification of the stage game. We call the latter objects "priors" to conform to past usage, but the language is inaccurate once we set the stage game in a repeated learning setting, and hence we occasionally emphasize this by referring to them as *stage-game* priors.

<sup>6</sup> We consider the case in which knowledge of opponents' play comes *only* from learning by observation and updating, and not from deduction based on opponents' rationality, so we do not require that players know their opponents' utility functions or beliefs. This is also the reason we do not need to consider moves by Nature that one player thinks that another thinks are possible, as no deductions are made, only learning from observations. Rubinstein and Wolinsky [1994], Battigalli and Guaitoli [1997] and Dekel, Fudenberg and Levine [1999] present solution concepts based on steady states in which players do make deductions based on rationality of the other players.

observation structures arising from play of a one-shot simultaneous-move game. Fudenberg and Kreps [1988, 1995] and Fudenberg and Levine [1993] examined learning, steady states, and self-confirming equilibria in games with non-trivial extensive forms.

We will not formally model the dynamics of learning, but will appeal informally to the idea that a steady state of a belief-based learning process must be a self-confirming equilibrium (Fudenberg and Levine [1993]). Thus, our focus is on how the information that players observe at the end of each round of play determines the set of self-confirming equilibria, and how these equilibria relate to the Nash equilibria of the game.

The key components of self-confirming (and Nash) equilibrium are each player  $i$ 's *beliefs* about Nature's move, her *strategy*, and her *conjecture* about the strategies used by her opponents. Player  $i$ 's beliefs, denoted by  $\hat{\mu}^i$ , are a point in the space  $\Delta(\Theta)$  of distributions over Nature's move, and her strategy is a map  $\sigma_i: \Theta_i \rightarrow \Delta(A_i)$ . The space of all such strategies is denoted  $\Sigma_i$ , and the player's conjectures about opponents' play are assumed to be a  $\hat{\sigma}_{-i} \in \times_{-i} \Sigma_{-i}$ , that is, a strategy profile of  $i$ 's opponents. The notation  $\hat{\mu}^i(\cdot | \theta_i)$  refers to the conditional distribution corresponding to  $\hat{\mu}^i$  and  $\theta_i$ , while  $\sigma_i(a_i | \theta_i)$  denotes the probability that  $\sigma_i(\theta_i)$  assigns to  $a_i$ .

**Definition:** A strategy profile  $\sigma$  is a self-confirming equilibrium with conjectures  $\hat{\sigma}_{-i}$  and beliefs  $\hat{\mu}_i$  if for each player  $i$ ,

$$(i) \quad \text{for all } \theta_i \text{ with } p(\theta_i) = 0, \hat{\mu}_i(\theta_i) = p(\theta_i)$$

and for any pair  $\theta_i, \hat{a}_i$  such that  $\hat{\mu}^i(\theta_i) \cdot \sigma_i(\hat{a}_i | \theta_i) > 0$  both the following conditions are satisfied

$$(ii) \quad \hat{a}_i \in \arg \max_{a_i} \sum_{a_{-i}, \theta_{-i}} u_i(a_i, a_{-i}, \theta_i, \theta_{-i}) \hat{\mu}^i(\theta_{-i} | \theta_i) \hat{\sigma}_{-i}(a_{-i} | \theta_{-i}),$$

and for any  $\bar{y}_i$  in the range of  $y_i$

$$(iii) \quad \begin{aligned} & \sum_{\{a_{-i}, \theta_{-i}: y_i(\hat{a}_i, a_{-i}, \theta_i, \theta_{-i}) = \bar{y}_i\}} \hat{\mu}^i(\theta_{-i} | \theta_i) \hat{\sigma}_{-i}(a_{-i} | \theta_{-i}) \\ &= \sum_{\{a_{-i}, \theta_{-i}: y_i(\hat{a}_i, a_{-i}, \theta_i, \theta_{-i}) = \bar{y}_i\}} p(\theta_{-i} | \theta_i) \sigma_{-i}(a_{-i} | \theta_{-i}). \end{aligned}$$



We say that  $\sigma$  is a *self-confirming equilibrium* if there is some collection  $(\hat{\mu}^i, \hat{\sigma}_{-i})_{i \in I}$  such that (i), (ii) and (iii) are satisfied.<sup>7</sup>

Condition (i) is a consequence of the assumptions that players observe their own types and that the types are i.i.d. over time; it will not apply in the setting considered in Section 5, where a player's type may be fixed once and for all. Condition (ii) says that any action played by a type of player  $i$  that has positive probability is a best response to her conjecture about opponents' play and beliefs about Nature's move. Condition (iii) says that the distribution of signals (conditional on type) that the player expects to see equals the actual distribution. This captures the least amount of information that we would expect to arise as the steady state of a learning process.

Note that the set of self-confirming equilibria can in general depend on the set  $\Theta$  of types that players think are possible, as this set determines the domain on which the beliefs  $\hat{\mu}_i$  are defined. On the other hand, the set of self-confirming equilibria does not depend on the exogenous stage-game priors  $\mu$ . To see why, note that a complete belief-based learning model would specify priors over both Nature's probability distribution and opponents' strategies. These priors would be updated over time, so that the steady state belief-conjecture pair  $(\hat{\mu}, \hat{\sigma}_{-i})$  need not be the same as the stage-game priors. In the learning process, different priors can lead to a different distribution over steady states; in our definition the set of self-confirming equilibria corresponds to the set of possible steady states for all initial conditions of the learning process.

We will sometimes consider the restriction of self-confirming equilibria to the case where players' beliefs about Nature satisfy certain restrictions. In particular, we say that a self-confirming equilibrium has “independent beliefs” if for all players  $i$  the beliefs  $\hat{\mu}^i$  are a product measure. Because the domain of  $\hat{\mu}^i$  is all of  $\Theta_0 \times \Theta_1 \times \dots \times \Theta_I$ , independence implies that player  $i$ 's beliefs about the types of her opponents do not depend on her own type. This restriction is most easily motivated in games where the

---

<sup>7</sup> It is appropriate to have a single  $\hat{\sigma}_{-i}$  for each player  $i$  in the definition because we assume that there is a single agent in each player role. This is called the “unitary” version of self-confirming equilibria; when we consider large populations and matching in Section 4 we allow for heterogeneous beliefs.

Note that  $i$ 's beliefs about opponents' play take the form of a strategy profile as opposed to a probability distribution over strategy profiles. The complications that arise due to correlations in conjectures are discussed in Fudenberg and Kreps [1988] and Fudenberg and Levine [1993]; we simplify by ignoring them here. Given this restriction, there is no further loss of generality in taking beliefs to be point conjectures. Battigalli [1987] defined a similar concept to the one above, as did Kalai and Lehrer [1993].

true distribution  $p$  is a product measure, that is, players' types are in fact independent, as in this case assuming independent beliefs amounts to saying that players understand this particular fact about the structure of the game. The following game demonstrates the effect of assuming independent beliefs.

*Example 1: Independent Beliefs*

Consider the following “bandit problem,” a one-person, two-type, two-action game, with two different states in  $\Theta_0$ . The actions are labeled *In* and *Out*; the player's types are labeled “*Timid*” ( $T$ ) and “*Brave*” ( $B$ ), the “Nature's type” states,  $\Theta_0$ , are labeled  $L$  and  $R$ . Both types get a payoff of 0 from *Out*. Payoffs from *In* are given in the table below.

	$L$	$R$
<i>Brave</i>	1	2
<i>Timid</i>	2	-1

Notice that *In* is a dominant strategy for the Brave type. Suppose the player does not observe Nature's move but does observe her own payoff. Suppose also that the objective distribution  $p$  on Nature's move assigns equal probability to the four states  $(B, L)$ ,  $(B, R)$ ,  $(T, L)$  and  $(T, R)$ . The Brave type has *In* as a dominant strategy, and so *Brave* will go *In* in every self-confirming equilibrium. Thus, since the player observes her payoff, the player learns the distribution of Nature's move conditional on *Brave*, so the only self-confirming equilibrium with independent beliefs has  $\hat{\mu} = p$  and both types playing *In*. However, there is also a self-confirming equilibrium without independent beliefs where the *Timid* type stays *Out* because the player believes that Nature plays  $R$  whenever the player is *Timid*, that is  $\hat{\mu}^i(R | T) = 1$ .

If there were only a *Timid* type, then clearly there would be a self-confirming equilibrium in which the player stays *Out*; this is exactly Example 5 (and similar to Example 1) of Jackson and Kalai [1997].

■

We explore the consequence of independent beliefs further in Example 7 of Section 5. This independence is one example of a situation where players use *a priori* information

about the joint distribution of types to restrict their beliefs. Another example is the case where players know the conditional distribution of Nature's type conditional on their own. Since players learn the distribution of their own type, they will be able to deduce the distribution of Nature's type,  $\theta_0$ . Thus Nash equilibria will only be self-confirming if players have correct beliefs about Nature's type. A third example is where players know the distribution of their opponents' types conditional on their own, but not necessarily that of Nature. In this case players will learn the distribution of players' types, that is, in any self-confirming equilibrium the marginal of  $\hat{\mu}$  on players' types will coincide with the marginal of  $p$  on players' types; this is one of the hypotheses of Proposition 6 below.

We are interested in the relationship between the set of self-confirming equilibria and the set of Nash equilibria. In a *Nash equilibrium*, each player's strategy must maximize her expected payoff given her stage-game prior about the distribution of  $\theta$  and correct conjectures about the play of the opponents.

**Definition:** A strategy profile  $\sigma$  is a Nash equilibrium with conjecture  $\hat{\sigma}_{-i}$  and beliefs  $\hat{\mu}_i$  if for each player  $i$ , and for any pair  $\theta_i, \hat{a}_i$  such that  $\hat{\mu}^i(\theta_i) \cdot \sigma_i(\hat{a}_i) > 0$

$$(ii) \quad \hat{a}_i \in \arg \max_{a_i} \sum_{a_{-i}, \theta_{-i}} u_i(a_i, a_{-i}, \theta_i, \theta_{-i}) \hat{\mu}^i(\theta_{-i} | \theta_i) \hat{\sigma}_{-i}(a_{-i} | \theta_{-i}),$$

and

$$(iii') \quad \hat{\sigma}_{-i} = \sigma_{-i}, \quad \hat{\mu}^i = \mu^i.$$

Note that (iii') has the further implication that  $\mu^i(\theta_i) > 0$  implies  $\hat{\mu}^i(\theta_i) > 0$ .<sup>8</sup> Note also that the definition of Nash equilibrium, unlike that of SCE, is unrelated to the observation structure; that is, the maps  $y_i(a, \theta)$  do not appear in the definition.

When the stage-game priors are diverse, we say that the Nash equilibrium has *diverse priors*. Finally, to distinguish the case where the beliefs are correct, that is  $\mu^i = p$  for all  $i$ , we say this is a *Nash equilibrium with correct priors*.

---

<sup>8</sup> This definition of Nash equilibrium allows for a player to believe that an opponent is not optimizing, since  $j$  can assign strictly positive probability to a type of  $i$  to which  $i$  assigns zero probability. To deal with this issue we could state the primitives of the game as conditional probabilities  $\mu^i(\theta_{-i} | \theta_i)$  and impose interim optimality even for own types to which one assigns zero probability. We chose to avoid this extra complexity in the notation.

The next result shows the restrictive effect that the assumption of private values has on the extent to which allowing for diverse priors increases the set of Nash equilibrium profiles. Later results in the paper show how the private-values assumption enables sharp comparisons of the sets of Nash and self-confirming equilibria.

**Proposition 0:** *The set of Nash equilibria of a two-player game with private values and arbitrary, possibly diverse, priors depends only on the support of the priors, and equals the set of Nash equilibria of the same game with common priors and the same support. That is, if  $\Theta$  is the fixed support of the distributions, the set of  $\sigma$  that are Nash equilibria for some  $\mu_1, \mu_2$  with support  $\Theta$  equals the set of  $\sigma$  that are Nash equilibria for some common prior  $\mu$  on  $\Theta$ .*

**Proof:** A player's prior about her own type does not matter to him, and so there is no harm in modifying it to reflect her opponent's belief. In a similar vein, neither player cares about Nature's type  $\theta_0$ . Thus, if  $\sigma$  is a Nash equilibrium where player 1's prior on  $\Theta_1 \times \Theta_2$  is  $\mu^1 = \mu_1^1 \times \mu_2^1$ , and two's prior is  $\mu^2 = \mu_1^2 \times \mu_2^2$ , then it is also an equilibrium when the priors are both  $\mu_1^2 \times \mu_2^1$ . ■

The following proposition summarizes how opposite extreme assumptions on players' observations effects the beliefs and strategy profiles in the self-confirming equilibria, leading in particular to ruling out profiles that are Nash equilibria only with diverse priors at one extreme, and allowing for a very large set of outcomes at the other. This leads us in the next section to explore the extent to which intermediate assumptions on observability can result in a closer relationship between Nash and self-confirming equilibria.

**Proposition 1:** *If players observe Nature's move, then in any self-confirming equilibrium the beliefs equal the objective distribution ( $\hat{\mu}_i = p$ ). Conversely, if players observe nothing ( $y_i(a, \theta) = \bar{y}$  for all  $a$  and  $\theta$ ) then the set of self-confirming equilibria allows for any beliefs  $\hat{\mu}$ , including  $\hat{\mu} = \mu$ , and includes all profiles of ex-ante undominated strategies.*

**Proof:** The proof of this statement, like those of most of the results in this paper, is an immediate consequence of the definitions, and hence often omitted. In this case, the conclusion follows from condition (iii) of the definition of SCE, which requires that the beliefs and conjectures generate the distribution of signals that the player actually sees. In particular, the first hypothesis implies that any profile  $\theta_{-i}$  generates a unique signal, so for each such profile  $\sum_{a_{-i}} \hat{\mu}^i(\theta_{-i} | \theta_i) \hat{\sigma}_{-i}(a_{-i} | \theta_{-i}) = \sum_{a_{-i}} p(\theta_{-i} | \theta_i) \sigma_{-i}(a_{-i} | \theta_{-i})$  hence  $\hat{\mu}^i(\theta_{-i} | \theta_i) = p(\theta_{-i} | \theta_i)$  as claimed. Regarding the second claim, the strategies are *ex-ante* undominated because there is only one agent in each player role, so that an agent's conjectures about the other players' strategies must be the same regardless of that agent's action and type, and the belief about Nature must also be conditionally independent of the action chosen given the type. ■

### 3. The Relationship Between Self-Confirming Equilibria and Nash Equilibria

In this section we focus on the relationship between self-confirming equilibria and Nash equilibria. Specifically we use a series of examples to explore the assumptions about observability under which the set of self-confirming equilibrium profiles with beliefs  $\hat{\mu}$  that are equal to some given  $\mu$  coincide with the set of Nash equilibrium profiles of the game where players' exogenous stage-game priors regarding Nature are this same value of  $\mu$ . We refer to this equality of profiles by saying that the sets of Nash and self-confirming equilibria with beliefs  $\mu$  coincide.

#### 3.1 The tension between Nash and self-confirming equilibria

As mentioned above, if players cannot observe or deduce their opponents' actions at the end of each period, then in general there can be self-confirming equilibria that are not Nash equilibria. So we begin by considering the case in which players either directly observe, or indirectly deduce from other observations, the realized actions of their opponents after each play of the game.

Let  $\theta_{-i} = (\theta_0, \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_I)$  denote a specification of the types of Nature and of all players other than  $i$ . Suppose that players observe their own utility, so that  $u_i(a, \theta) \neq u_i((a_i, a'_{-i}), (\theta_i, \theta'_{-i}))$  implies  $y_i(a, \theta) \neq y_i((a_i, a'_{-i}), (\theta_i, \theta'_{-i}))$ . With generic payoffs the map  $u_i(a, \theta)$  is 1-1, and both the actions of other players,  $a_{-i}$ , and Nature's move,  $\theta$ , can be uniquely determined from  $y_i$ . Consequently, the only beliefs and conjectures that are self confirming are the correct ones.

Next suppose that  $y_i = a_{-i}$ , that is, players observe their opponents' actions. Then in a self-confirming equilibrium players must know the conditional distribution of opponents' actions given their own type. Suppose in addition that the game is a game of private values, that is,  $u_i(a, \theta) = u_i(a, \theta_i)$ . Since a player's payoffs do not depend on her opponents' types, in a game with private values, any strategy for player  $i$  that is a best response to conjectures and beliefs consistent with the observed distribution over actions must also be a best response to the true distributions of opponents' actions and Nature's move. This implies that with private values and observed actions every self-confirming equilibrium has the same strategies as a Nash equilibrium of the game with the correct and hence common priors.

**Proposition 2:** *If either*

*(i) payoffs are generic ( $u_i(a, \theta) \neq u_i(a', \theta')$  if either  $a \neq a'$  or  $\theta \neq \theta'$ ) and observed, or*

*(ii) there are private values and observed actions,*

*then the set of strategy profiles in self-confirming equilibria coincides with the set of Nash equilibrium profiles of the game with the correct (hence common) prior.*

Thus under either condition of the proposition, if the stage-game priors in a given Bayesian game are not common, and in addition, if the set of Nash equilibria of that game differs from the set of Nash equilibria with the correct prior (that is if the presumption of diverse priors has any significance), then the Nash equilibria of the game with diverse priors will not coincide with the self-confirming equilibria. This is demonstrated in Example 2 below.

Jackson and Kalai [1997], Theorem 1, prove a related result. They assume that payoffs are “privately observable,” and that “social learning implies private learning,” and reach the similar conclusion that Bayesian equilibria of a “recurring” game converge to Bayesian equilibria of the stage game (with correct priors).<sup>9</sup> Their Example 2, of a repeated private-values first-price sealed-bid auction with observed winning bids is an instance of their result, and of ours. Proposition 2 (i)’s assumption of observed payoffs implies that payoffs are privately observed, but this is not sufficient for the result, as shown by Example 4. The assumptions of Proposition 2 (ii) also imply that payoffs are privately observed; but as before the latter is not sufficient for our conclusion.

*Example 2: Nash Equilibria that are Not Self-Confirming Equilibria*

We consider a game with a column player,  $C$ , and two row players,  $R1$  and  $R2$ . Nature chooses  $L$  or  $R$ , with equal probability; the column player observes Nature’s choice of  $L$  or  $R$ , while the two other players do not. Thus players  $R1$  and  $R2$  each have a single type, player  $C$  has two types,  $L$  and  $R$ , and the set  $\Theta_0$  of Nature’s types is empty.

In this game,  $C$ ’s payoff depends only on her own action and type, but not on the actions taken by the row players: specifically,  $C$ ’s actions are labeled  $l$  and  $r$ , and  $C$  gets 1 for choosing the same as Nature, and 0 for choosing the opposite. The row players’ payoffs each depend on the column player’s action and their own action, as shown in the following two matrices.

---

<sup>9</sup> A game has *privately observed payoffs* if for all players  $i$  and all  $a, \theta, \bar{a}_{-i}, \bar{\theta}_{-i}$ ,  $y_i(a, \theta) = y_i((a_i, \bar{a}_{-i}), (\theta_i, \bar{\theta}_{-i}))$  implies  $u_i(a, \theta) = u_i((a_i, \bar{a}_{-i}), (\theta_i, \bar{\theta}_{-i}))$ . Jackson and Kalai consider *recurring games* in which successive players play a stage game, and observe a public signal, but need not know the distribution of their own type or actions played by former players in the same role. Nature picks a stage game distribution from a fixed distribution, as in our discussion of exchangeability in Section 5c, and players have a correct and hence common prior about this initial move by Nature. Their condition that *social learning implies private learning* says, roughly, that whenever the distribution of past public signals permits an observer to forecast as well as if he knew the distribution of Nature’s move, then the players’ forecasts conditional on past public signals, their own type, and own action, are the same as if they also knew the distribution of Nature’s move. This is a condition on the equilibrium of the overall recurring game; it is always satisfied when actions are observable and types are independently distributed.

$RI$	$L$	$R$
$U$	$\frac{3}{4}$	$\frac{3}{4}$
$D$	1	0

$R2$	$l$	$r$
$U$	$\frac{3}{4}$	$\frac{3}{4}$
$D$	0	1

This is a game with private values, because the row players' payoffs depend only on the column player's action, not her type. We will assume that actions are observed at the end of each round of play in the learning environment, so that Proposition 2 applies. Clearly the column player has a dominant strategy of playing  $l$  when type  $L$  and  $r$  when type  $R$ , so in a self-confirming equilibrium, the column player plays  $l$  on  $L$  and  $r$  on  $R$ , and so the column player takes each action half of the time. The row players observe this, so must play  $U$ .

Now suppose that  $RI$ 's stage-game prior assigns probability .9 to type  $L$  and .1 to  $R$ , while  $R2$ 's stage-game prior is the reverse, with .1 probability of type  $L$  and .9 to  $R$ . In a Nash equilibrium,  $C$  plays  $l$  upon observing  $L$  and  $r$  upon observing  $R$ , and the row players know this. Given the stage-game priors, this implies that  $RI$  and  $R2$  believe that they will face the actions  $l$  and  $r$  respectively .9 of the time. Consequently, in this Nash equilibrium with diverse stage-game priors,  $RI$  and  $R2$  will both choose  $D$ . However, this is not a Nash equilibrium for any common prior, and so it is not a self-confirming equilibrium for any  $p$  when the column player's action is observed.

We see in this example that, when players observe actions, the self-confirming equilibria in which beliefs are equal to the stage-game priors is unique, and is different from the Nash equilibrium. When players observe nothing at all, the set of self-confirming equilibria with beliefs equal to the stage-game priors includes the Nash equilibrium, but in fact imposes no restrictions at all on the play by  $RI$  and  $R2$  since the row players will not know anything about column's choice.

One way of summarizing this example is to say that although Nash equilibrium requires two players to agree about an opponent's *strategy*, these players can have different forecasts about the distribution of that opponent's *actions* if they have different beliefs about the distribution of that player's type. In contrast, with observed actions players correctly forecast the distribution of opponent's actions in any self-confirming



equilibrium, but they can have different beliefs about the distribution of Nature's move and about the opponent's strategy

■

As we noted earlier, the set of self-confirming equilibria can depend on the set  $\Theta$  of types that the players think are possible, as this set is the support of the allowed beliefs, and (as in the second equilibrium discussed in Example 1) some beliefs about the consequences of an action might keep a player from playing it. However, when players know their payoffs as a function of the strategy profile, the game is one of private values, and as in Proposition 2b, each player's best response depends only on his forecast of the distribution of his opponents' actions. This leads to the following conclusions:

**Proposition 3:**

*(i) When there are private values and the distribution  $p$  of Nature's move assigns probability 1 to a single profile of types, the definition of self-confirming equilibrium in this paper is equivalent (up to the different assumptions on the observation structure) to the definition given in Fudenberg and Levine [1993] for the complete-information case, and in particular is independent of the set  $\Theta$ .*

*(ii) If in addition actions are observed each round, the set of self-confirming equilibria reduces to the set of Nash equilibria of the complete-information game corresponding to the realized payoff functions.*

**Proof:** For part (i), recall that with private values, each player's set of best responses depends only on the player's conjecture about the play of the opponents. Thus the players' beliefs about Nature's move are irrelevant for whether a strategy profile is a self-confirming equilibrium; all that matters is that each player's strategy is a best response to the player's conjectures about opponents' play, and that these conjectures are consistent with the distribution of signals that the player observes; this is exactly the definition of self-confirming equilibrium in Fudenberg and Levine [1993], except that they specialize to the case where the signal observed is the terminal node of an extensive-form game.

Part\_(ii) follows because, as in to the proof of Proposition 1 condition (iii) of the definition of SCE requires conjectures to be correct when actions are observed.

■

In Section 4, we develop a small extension of this proposition that applies to a common design for game theory experiments.

The next two examples explore the role of observing actions in the relationship between Nash and self-confirming equilibria in games that do not have private values. We summarize their implications as follows.

**Proposition 4 :** *(i) Without private values ( $u_i(a, \theta) \neq u_i(a, \theta_i)$  for some  $(a, \theta)$ ), if neither types nor payoffs are observed, but actions are ( $y_i(a, \theta) = a$ ), there can be self-confirming equilibria with correct beliefs about Nature ( $\hat{\mu}^i = p$ ) that are not Nash even with correct priors (Example 3).*

*(ii) Even if the set of strategy profiles in self-confirming equilibria with beliefs  $\mu = \hat{\mu}$  coincides with the set of Nash equilibria, conjectures about opponents' play may fail to be correct ( $\hat{\sigma}_{-i} \neq \sigma_{-i}$ ). Consequently the profile can fail to be self confirming once actions are added to the available information (Example 4).*

*Example 3: Self-confirming equilibria that are not Nash with observed actions*

Player  $R$  and player  $C$  each choose either  $-1$  or  $1$ . Player  $R$ 's type is either  $+1$  (with probability  $2/3$ ) or  $-1$  (with probability  $1/3$ ), and player  $R$ 's payoff is her action times her type, so player  $R$  plays  $+1$  when type  $1$  and  $-1$  when type  $-1$ . Player  $C$ 's payoff is the product of player  $R$ 's type and the two actions, so the unique Nash equilibrium with the correct prior has player  $C$  play  $+1$ . If all that player  $C$  observes is player  $R$ 's action, then player  $C$  can have correct beliefs about Nature's move and conjecture that player  $R$  plays  $+1$  when type  $-1$  and mixes  $1/2$ -  $1/2$  when type  $+1$ . In this case the best response is for player  $C$  to play  $-1$ . Consequently, player  $C$  plays  $-1$  in this self-confirming equilibrium.

■

*Example 4: A game where when payoffs are observed, Nash equilibrium and self-confirming equilibrium are equivalent iff actions are not observed.*

Consider a two-player game in which Nature chooses the left or right matrix. Neither player has private information. Proposition 2 (i) does not apply because the payoffs include ties as shown below.

	$A$	$B$
$A$	1, 1	0, 0
$B$	0, 0	0, 0

	$A$	$B$
$A$	0, 0	1, 1
$B$	1, 1	0, 0

To analyze Nash equilibria, suppose that the stage game prior is that both players think the left matrix is chosen with probability  $1 - \varepsilon$ . The strategic form for this game given the common beliefs  $\mu$  is

	$A$	$B$
$A$	$1 - \varepsilon, 1 - \varepsilon$	$\varepsilon, \varepsilon$
$B$	$\varepsilon, \varepsilon$	0, 0

The unique Nash equilibrium for the specified beliefs is  $(A, A)$ .

Now suppose that in the learning environment, the true probability of the left matrix is  $\varepsilon$ . If players observe only their payoffs, then  $(A, A)$  is a self-confirming equilibrium with beliefs  $(1 - \varepsilon, \varepsilon)$  and conjecture that the opponent is playing  $B$ : in this case each player believes that playing  $A$  yields 1 with probability  $\varepsilon$ , and  $B$  yields 0. However, if players were to also observe actions, then the Nash equilibrium  $(A, A)$  would no longer be self confirming.

■

### 3.2 Examples where Nash equilibria and self-confirming equilibria do coincide.

Proposition 2 shows that with private values and observed actions, the set of strategy profiles in self-confirming equilibria coincides with the set of Nash equilibrium profiles of the game with the correct (hence common) prior, and Example 2 shows that on these conditions there can also be Nash equilibria with diverse priors that are not self-confirming with respect to any beliefs. Our next example shows that the reverse conclusion can hold: the sets of Nash equilibria and self-confirming equilibria for a given diverse prior  $\mu$  can coincide, even when actions are observed.

**Proposition 5:** *The Nash equilibria with diverse priors  $\mu$  and the self-confirming equilibria with beliefs  $\hat{\mu} = \mu$  may coincide, whether or not players observe actions.*

This is demonstrated by the following example:

*Example 5: A game where Nash equilibrium and self-confirming equilibrium coincide for a specific diverse belief about Nature's move*

	$L$	$R$
$U$	1, 1	0, 0
$D$	0, 0	-1, -1

	$L$	$R$
$U$	-1, -1	0, 0
$D$	0, 0	1, 1

This is a two-player game in which Nature chooses the left ( $l$ ) or right ( $r$ ) payoffs, and neither player observes Nature's move. The row player believes the left payoffs are chosen, the column player believes the opposite:  $\mu^1(l) = \mu^2(r) = 1$ . So the unique Nash equilibrium is for the row player to play  $U$  and the column player  $R$ , with payoffs (0, 0). Whether or not players observe their opponent's actions or their own utility, this profile is self-confirming with beliefs equal to the given stage-game priors. However, the subset of self-confirming equilibria with beliefs in which  $\hat{\mu}^1 = \hat{\mu}^2$  is either  $(U, L)$ ,  $(D, R)$ , or the entire strategy space.

■

Note that the diverse priors in the example are significant: the set of Nash (and self-confirming) equilibria with the diverse priors differs from the set of Nash (and self-confirming) equilibria with a common prior. The example demonstrates this point using *ex-ante* dominating strategies, in which it is irrelevant what players observe; in an example in the appendix the players do care about their opponents' actions, and in that example the sets of pure-strategy Nash and self-confirming equilibria with a specific diverse prior coincide if players observe either their own payoffs, or the opponents' actions.

To summarize, we have seen that observing actions is neither necessary nor sufficient for self-confirming and Nash equilibria to coincide. Moreover, a Nash equilibrium with a diverse prior that is a self-confirming equilibrium when players only observe their payoff need no longer be a self-confirming equilibrium if players can observe actions as well as payoffs. This suggests that, loosely speaking, the “best” case for Nash and self-confirming equilibria to coincide for some diverse beliefs  $\mu$  is when actions are observable and payoffs are not. Moreover, Proposition 2 shows that with observed actions and private values, any self-confirming equilibrium is a Nash equilibrium with the correct beliefs. This suggests that the simplest case where the two sets might be equal with diverse beliefs is when actions are observed and payoffs depend on Nature's move as well as own type, that is  $u_i(a, \theta) = u_i(a, \theta_i, \theta_0)$ . But even in this case, the equivalence of NE and SCE is not guaranteed: In example 2 we saw that if players observe actions and not payoffs, while each player can correctly infer the opponents' distribution of actions as a function the player's own type, this need not force them to agree about the distribution of private types, and it also does not force them to agree about  $\theta_0$ , the portion of Nature's move that is unknown to everyone.

**Proposition 6:** *If actions are observed but payoffs are not, ( $y_i(a, \theta) = a$ ), the set of Nash equilibria coincides with the set of self-confirming equilibria under the additional hypothesis that the marginal of  $\mu$  on players' types coincides with the marginal of  $p$  on players' types.*

**Proof:** Since actions are observed, in a self-confirming equilibrium each player learns the distribution of opponent actions conditional on his own type and plays a best response. In a Nash equilibrium each player correctly knows opponent strategies, and infers the

distribution of actions conditional on his own type. Since the marginal  $\mu$  is the same as  $p$  the distribution of actions conditional on own type must be the same as in the self-confirming case. ■

#### 4. Large Populations and Heterogeneous Beliefs

Next we focus on a class of games of special interest in learning theory and in experimental economics: games in which players are randomly matched to play a “stage” game. In this setting it is natural to think of  $p$  as the distribution of types for a given match, but we must also consider the relationship between the matching process and the  $p$  from which players draw their observations.

Suppose that individuals in a given player role are independently matched in each period with opponents in other roles, and that after each match is made Nature draws types for that match according to the distribution  $p$ . We know from Fudenberg-Levine [1993] that when there are multiple agents in each player role, there can be “heterogeneous” self-confirming equilibria in which different agents in the same role play different strategies and have different conjectures. Thus, when types are chosen independently over time, and separately for each match, the appropriate definition allows the beliefs and conjectures of the agents to vary with the strategy chosen.

**Definition:** A strategy profile  $\sigma$  is a heterogeneous self-confirming equilibrium if for each player  $i$  there exists a finite set of strategies  $\{\sigma_i^k : k \in K\} \subset \Sigma_i$  such that  $\sigma_i$  is in the convex hull of  $\{\sigma_i^k : k \in K\}$  and such that for each  $\sigma_i^k$  there are conjectures  $\hat{\sigma}_{-i}$  and beliefs  $\hat{\mu}_i$  (both of which can depend on  $\sigma_i^k$ ), such that

$$(i) \quad p(\theta_i) = \hat{\mu}_i(\theta_i),$$

and for any pair  $\theta_i, \hat{a}_i$  such that  $\hat{\mu}_i^i(\theta_i) \cdot \sigma_i(\hat{a}_i | \theta_i) > 0$  both the following two conditions are satisfied

$$(ii) \quad \hat{a}_i \in \arg \max_{a_i} \sum_{a_{-i}, \theta_{-i}} u_i(a_i, a_{-i}, \theta_i, \theta_{-i}) \hat{\mu}_i^i(\theta_{-i} | \theta_i) \hat{\sigma}_{-i}(a_{-i} | \theta_{-i}),$$

and for any  $\bar{y}_i$  in the range of  $y_i$

$$\begin{aligned}
\text{(iii)} \quad & \sum_{\{a_{-i}, \theta_{-i} : y_i(\hat{a}_i, a_{-i}, \theta_i, \theta_{-i}) = \bar{y}_i\}} \hat{\mu}^i(\theta_{-i} \mid \theta_i) \hat{\sigma}_{-i}(a_{-i} \mid \theta_{-i}) \\
& = \sum_{\{a_{-i}, \theta_{-i} : y_i(\hat{a}_i, a_{-i}, \theta_i, \theta_{-i}) = \bar{y}_i\}} p(\theta_{-i} \mid \theta_i) \sigma_{-i}(a_{-i} \mid \theta_{-i}).
\end{aligned}$$

With this definition,  $\sigma_i$  corresponds to the aggregate play of the population of player  $i$ 's; each individual in the role of player  $i$  is playing one of the  $\sigma_i^k$ , and the individual's belief and conjecture depend on the strategy she plays.

Intuitively, allowing different agents in the role of player  $i$  to have different self-confirming beliefs makes no difference when the beliefs must be correct. For this reason, allowing for heterogeneous beliefs makes no difference when players observe actions and types at the end of each round, nor when players observe their realized payoffs and this information reveals the types and actions, nor when players observe their opponent's actions and there are private values. In all of these cases, as in Propositions 1 and 2, the heterogeneous self-confirming equilibria coincide with the Bayesian Nash equilibria with the correct beliefs.

A bandit problem provides an easy way illustrate the difference this heterogeneity can make. Consider a simplified version of the bandit problem in Example 1, where player one now has the single type "Timid." Playing *Out* gives this type a payoff of 0, playing *In* gives 2 if Nature plays  $L$  and -1 if Nature plays  $R$ , and the distribution of Nature's move is such that it is optimal for the player to always play *In*. The player knows the structure of the model, but does not know the distribution of Nature's move; the player observes her payoff but doesn't directly observe Nature's move. In this game, there is no self-confirming equilibrium where the player randomizes, since a player who plays *In* with positive probability must know the expected payoff to *In*. But there are heterogeneous self-confirming equilibria where some players play *In* and others stay *Out*: The players who play *In* are playing a best response to the true distribution of Nature's move; the players who play *Out* believe that Nature plays  $R$  with probability more than  $2/3$ , and never see evidence that this belief is mistaken.

Example B in the appendix shows that heterogeneity can make a difference even in games where any SCE with unitary beliefs has beliefs that equal the true distribution.

■

## 5. The Joint Distribution of Nature's Moves over Time, Agents, and Players

So far we have supposed that that Nature makes independent draws from  $p$  each period. In this section we consider alternative stochastic processes for the specification of types, and explore how this affects the notion of self-confirming equilibrium.

### 5a) Perfect Correlation: A single type profile in all matches and all dates

One specification that is often used in experiments is for Nature to make a once-and-for-all choice of a single profile of types  $\theta$  that will apply to all matches.<sup>10</sup> In this case, even if the profile is chosen according to some non-degenerate distribution  $p$ , this distribution is not directly relevant to the long-run outcome. Thus, the appropriate definition of self-confirming equilibrium replaces  $p$  in condition (iii) with the degenerate distribution that chooses  $\hat{\theta}$  with probability one. This is “appropriate” in the sense that the players learn only about the particular draw, and the relative probability of the types that did not occur does not influence the set of possible steady states.<sup>11</sup> Combining this observation with Proposition 3, we conclude that if Nature picks the payoff functions according to a possibly non-degenerate distribution, players know the map from actions to their own payoffs (i.e. private values), and actions are observed, the set of self-confirming equilibria reduces to the set of Nash equilibria of the complete-information game corresponding to the realized payoff functions. This is the case in all the experimental papers described in footnote 12.<sup>12</sup>

---

<sup>10</sup> Examples include Cox *et al.*, Chen (2000), Mitropoulos (2001) , and Oechssler and Schippers [2002].

<sup>11</sup> Battigalli and Guaitoli [1997 Section 1.4] study exclusively the unitary-beliefs version of this case where types are drawn once and for all. They provide a version of the corresponding form of self-confirming equilibrium, and use it to analyze an example which has the same property as example 7 below: their Proposition 1 and discussion on p. 116 imply that in one state the players' behavior coincides with their behavior where the state is commonly known, whereas in another state there is a self-confirming equilibrium in which players' behavior can rely on the state not being commonly known. Their footnote 4 states a special case of Proposition 2' below, and their footnote 10 discusses how to modify their definition to correspond to the type-heterogeneous notion defined at the end of this section; they do not explore the impact of this modification.

<sup>12</sup> Of these, only Cox *et. al.* told subjects about the distribution from which the fixed types were drawn; in the relevant treatments of their experiments, subjects were told nothing at all about the payoff functions of their opponents.



**5b) Fixed types for each agent, but diversity across agents in the same role**

Now we suppose that each agent's type is fixed once and for all, before any matches have taken place, and that players are randomly matched each period, with the matching process independent of the players' types. Then even if the distribution  $p$  from which the types were drawn exhibits correlation, the matching process will lead the distribution of types in each match to be independent, and the distribution that governs the learning process will be the product of the realized marginal distributions on each player's type. The heterogeneous self-confirming equilibrium defined in Section 4 is appropriate when Nature's move is i.i.d. over time, since a given agent eventually receives many observations of the distribution of signals corresponding to each possible type  $\theta_i$  in the support of  $p$ , but it is not appropriate in the present case where types are fixed once and for all, as each agent is only in the role of a single type, and there is no reason that beliefs across types should be consistent with updating from a common prior.<sup>13</sup> Therefore, instead of imposing that restriction, we allow each type  $\theta_i$  to have any "interim belief"  $\tilde{\mu}^{\theta_i}$  that is consistent with that type's observations. Similarly, when types are fixed, conjectures may depend on types. The following notion of type-heterogeneous self-confirming equilibrium captures the idea that types are fixed initially, but that players are subsequently matched with opponents whose types have been drawn according to  $p$ .

**Definition:** A strategy profile  $\sigma$  is a type-heterogeneous self-confirming equilibrium if for each player  $i$ , and for each  $\hat{a}_i$  and  $\theta_i$  such that  $p(\theta_i) \cdot \sigma_i(\hat{a}_i | \theta_i) > 0$  there are conjectures  $\hat{\sigma}_{-i}$  and interim beliefs  $\tilde{\mu}^{\theta_i}$  (both of which can depend on  $\hat{a}_i$  and  $\theta_i$ ), such that both the following conditions are satisfied

$$(ii) \quad \hat{a}_i \in \arg \max_{a_i} \sum_{a_{-i}, \theta_{-i}} u_i(a_i, a_{-i}, \theta_i, \theta_{-i}) \tilde{\mu}^{\theta_i}(\theta_{-i}) \hat{\sigma}_{-i}(a_{-i} | \theta_{-i}),$$

$$(iii) \quad \text{for any } \bar{y}_i \text{ in the range of } y_i$$

---

<sup>13</sup> If no restrictions are imposed on the prior, then *any* collection of interim beliefs  $(\tilde{\mu}^{\theta_i})_{\theta_i \in \Theta_i}$  can be generated from a prior  $\mu_i$  by setting  $\mu_i(\theta_{-i}, \theta_i) = \mu_i(\theta_i) \tilde{\mu}^{\theta_i}(\theta_{-i})$  for some marginals  $\mu_i(\theta_i)$ , but the interim definition allows for each type of player  $i$  to think all types are independently distributed while also allowing different types to have different beliefs.

$$\begin{aligned}
& \sum_{\{a_{-i}, \theta_{-i} : y_i(\hat{a}_i, a_{-i}, \theta_i, \theta_{-i}) = \bar{y}_i\}} \tilde{\mu}^{\theta_i}(\theta_{-i}) \hat{\sigma}_{-i}(a_{-i} \mid \theta_{-i}) \\
&= \sum_{\{a_{-i}, \theta_{-i} : y_i(\hat{a}_i, a_{-i}, \theta_i, \theta_{-i}) = \bar{y}_i\}} p(\theta_{-i} \mid \theta_i) \sigma_{-i}(a_{-i} \mid \theta_{-i}).
\end{aligned}$$

Notice that condition (i), which required correct beliefs about one's own types, is no longer appropriate, since each agent in a player role does not observe the distribution of types in that player role.

In the following variant on Example 1 “*In* when *Brave*, *Out* when *Timid*” is not a self-confirming equilibrium, and it is not a heterogeneous self-confirming equilibrium, but it is a type-heterogeneous self-confirming equilibrium.

*Example 6:* Consider a two-player game with two types of player 1. Player 1 can play *In* or *Out*; *Out* gives payoff 0, while the payoff to *In* depends on player 1's type and player 2's action as shown below. (Player 2's payoff is not shown, because we assume that for player 2 *L* is a dominant strategy.)

	<i>L</i>	<i>R</i>
<i>Brave</i>	1	2
<i>Timid</i>	2	-1

Suppose that player 1 observes her type, action, and realized payoff, but not the action of player 2. The only self-confirming equilibrium is for both types of player 1 to play *In*, and this is also the only heterogeneous self-confirming equilibrium: Since for the *Brave* type *In* is a dominant strategy, when types are i.i.d., any undominated strategy for an agent in the role of player 1 will cause the agent to learn that 2 plays *L*. However, “*Brave In*, *Timid Out*” is a type-heterogeneous self-confirming equilibrium, because a *Timid* type can stay *Out* forever and never learn the true distribution of player 2's play.<sup>14</sup>

---

<sup>14</sup> The difference between this example and a bandit problem such as Example 1, where player 2 is replaced by Nature, is that player 1 can think Nature's move is correlated with her type, but player 1's conjecture about player 2 must correspond to a strategy for player 2, and since player 2 does not observe player 1's type, player 2's strategy cannot depend on it.

■

In the examples so far, type-heterogeneity mattered because it let different types play different actions and maintain different beliefs. In the next example, type-heterogeneity allows types to play the same action.

*Example 7: Independent heterogeneous Self-Confirming Equilibria versus Independent type-heterogeneous Self-Confirming Equilibria*

Consider another variant of the one-player game of Example 1, where the payoffs for *Out* remain 0 and those for *In* are

	<i>L</i>	<i>R</i>
<i>Brave</i>	2	-1
<i>Timid</i>	-1	2

Here both types can stay *Out* only if they disagree about Nature's move: *Brave* must believe *R* and *Timid* must believe *L*. Suppose in fact that players observe nothing, so that behavior depends only on stage-game priors. If the players' types are drawn anew each period and beliefs are restricted to be independent, then in any self-confirming equilibrium  $\sigma(In|Brave) + \sigma(In|Timid) \geq \sigma(Out|Brave) + \sigma(Out|Timid)$  since the beliefs corresponding to any  $\sigma^k$  must lead them to play *In* either when they are *Brave* or when they are *Timid* (or both). On the other hand, if players' types are drawn once and for all, they can stay *Out* forever (each type can have constant beliefs justifying *Out*).

■

### 5c) More general stochastic structures

Underlying our notion of a steady state is the idea that players repeatedly sample from a fixed distribution that does not change over time. Suppose we consider the more general class of exchangeable processes for types, which have a representation as a “prior” probability distribution over (conditionally) i.i.d. processes, and for the time being suppose there is a single agent in the role of each player. Then we can think of Nature making a single once-and-for-all draw  $\hat{p}$  from the class of i.i.d. processes, and the “appropriate distribution” to use in the definition of a self-confirming equilibrium is the

$\hat{p}$  drawn by Nature; the fact that players “could have” faced some different distribution and that the overall distribution was  $p$  is not relevant in the long-run steady state.<sup>15</sup> Note that the exchangeable model nests both the case of a single once-and-for-all draw  $\hat{\theta}$  and the case where each period’s  $\theta$  is an independent draw from  $p$ . Note also that the distribution from which Nature chooses  $\hat{p}$  does influence the ex-ante distribution over steady states.

Thus, we can extend the discussion of private values.

**Proposition 2’:** *With private values ( $u_i(a, \theta) = u_i(a, \theta_i)$ ) and observed actions ( $y_i(a, \theta) = a$ ), a self-confirming equilibrium is a Nash equilibrium in a game with stage-game priors equal to the “realized distribution” of types,  $\hat{p}$ .*

One can also consider the class of ergodic processes instead of exchangeable ones. It is natural in that case to think of  $p$  as the invariant distribution. Notice in this case that players are not actually drawing from  $p$  each period, rather they are drawing from time-varying distributions which average out to  $p$ . If players believe that the true process is exchangeable, then beliefs in steady states will still satisfy the self-confirming conditions of Section 2 with respect to this ergodic distribution.<sup>16</sup>

---

<sup>15</sup> The relationship between the ex-ante distribution  $p$  and the realized value of  $\hat{p}$  here is analogous to the relationship between the prior  $\Gamma$  and the realized  $\tau$  in Jackson and Kalai [1997].

<sup>16</sup> Of course, sophisticated players might realize that Nature’s moves do not have an exchangeable distribution, in which case our definition of self-confirming equilibrium with  $p$  set equal to the ergodic distribution would not correspond to steady states.

## Appendix

### Example A

Since Example 5 involves dominant strategies, it is not very interesting from a game-theoretic perspective. The next, more complicated, example, due to Phil Reny, shows that dominant strategies are not required for the property that the *pure*-strategy Nash and self-confirming equilibria coincide if players observe actions or payoffs or both. However, the equivalence fails for mixed-strategy equilibria.

In this game there are three states of Nature  $\theta_0^I, \theta_0^{II}, \theta_0^{III}$  and no types. There are two players, a row and a column player; each chooses between three actions  $T, M, B$ . Payoffs in each of the states is given in the table below.

	$\theta_0^I$		
	$T$	$M$	$B$
$T$	0, 1	0, $\frac{1}{2}$	1, -1
$M$	5, 0	5, 5	$\frac{1}{2}$ , 5
$B$	0, 0	0, 5	-1, 0

	$\theta_0^{II}$		
	$T$	$M$	$B$
$T$	0, -1	0, $\frac{1}{2}$	-1, 1
$M$	5, 0	5, 5	$\frac{1}{2}$ , 0
$B$	0, 0	0, 5	1, 0

	$\theta_0^{III}$		
	$T$	$M$	$B$
$T$	0, 1	0, 5	1, 1
$M$	5, 0	5, 5	5, 0
$B$	0, 0	0, 5	1, 0

Beliefs about and the actual distribution of Nature's move are given below

	$\theta_0^I$	$\theta_0^{II}$	$\theta_0^{III}$
$\mu_1$	$1-2\epsilon$	$\epsilon$	$\epsilon$
$\mu_2$	$\epsilon$	$1-2\epsilon$	$\epsilon$
$\mu$	$\epsilon$	$\epsilon$	$1-2\epsilon$

To analyze the game, note that if 2 plays  $T$  or  $M$  it is a strict best response for 1 to play  $M$ ; if 1 plays  $M$  or  $B$  it is a strict best response for 2 to play  $M$ . Hence the relevant portion of the game involves 2 playing  $B$  or 1 playing  $T$ . Payoffs in these cases are summarized below.

	$\theta_0'$	$\theta_0''$	$\theta'$
$u_1(T, B, \theta_0) = u_2(T, T, \theta_0) =$	1	-1	1
$u_1(M, B, \theta_0) = u_2(T, M, \theta_0) =$	$\frac{1}{2}$	$\frac{1}{2}$	5
$u_1(B, B, \theta_0) = u_2(T, B, \theta_0) =$	-1	1	1

With the given stage-game priors the pure-strategy Nash equilibria are  $(M, M)$  and  $(T, B)$ . The latter is not a Nash equilibrium with a common prior. If players observe payoffs, or opponents' actions, then there are two pure strategy self-confirming equilibria with beliefs  $\mu_1, \mu_2$ : one with the strategy profiles  $(M, M)$  and the other with  $(T, B)$ .

To see that this equivalence fails for mixed strategies note first that for  $\varepsilon$  small there is a mixed-strategy Nash equilibrium in which column plays  $B$  with probability  $10/(11-13\varepsilon)$  and  $M$  with complementary probability; and row plays  $T$  and  $M$  with these probabilities. If payoffs and opponents' actions are observed and players play these strategies then the players would learn Nature's distribution, so this is not a self-confirming equilibrium.

#### Example B

This example shows that heterogeneity can make a difference even in games where any SCE with unitary beliefs has beliefs that equal the true distribution.

This is a two-player game without types (equivalently, the type space  $\Theta$  consists of a single point.) Players observe payoffs, but not their opponents' actions.

	$LL$	$L$	$C$	$R$	$RR$
$U$	1,0	0,0	1,1	0.01,2	0.03,-12
$M$	1,0	1,0	2,1	0.5,-10	0.6,-11
$D$	0,0	1,0	1,1	0.04,-13	0.02,2

The row player plays  $U$  only if he believes column plays  $LL$  with probability 1, and  $D$  only if he believes column plays  $L$  with probability 1. So in a unitary SCE, row cannot play both  $U$  and  $D$ . If and only if row plays  $U$  with sufficiently high probability

will column play  $R$ , which will be known to row (by observing payoffs), and then column cannot play  $U$ . Similarly  $RR$  is ruled out, and both  $LL$  and  $L$  are strictly dominated, so the unique unitary SCE is the unique NE (namely  $M, C$ ). However, there is a heterogeneous self-confirming equilibrium in which some row agents play  $U$  believing  $LL$ , other row agents play  $D$  believing  $L$ , the two kind of agents are selected to play with equal probability, and column plays  $C$ .

## References

- Banerjee, Abhijit and Rohini Somanathan [2001] "A Simple Model of Voice," *Quarterly Journal of Economics*; 116 (1), 189-228.
- Battigalli, Pierpaolo [1987] "Comportamento Razionale Ed Equilibrio Nei Giochi E Nelle Situazioni Sociali,," unpublished undergraduate dissertation, Bocconi University, Milano.
- Battigalli, Pierpaolo and D. Guaitoli [1997] "Conjectural Equilibria and Rationalizability in a Game with Incomplete Information," in *Decision, Games and Markets*, P. Battigalli, A. Montesano and F. Panunzi, Eds., Dordrecht: Kluwer Academic Publishers.
- Cox, James, Jason Shachat, and Mark Walker [2001] "An Experiment to Evaluate Bayesian Learning of Nash Equilibrium Play," *Games and Economic Behavior* 34(1), 11-33.
- Chen, Yan [2000] "An Experimental Study of Serial and Average Cost Pricing Mechanisms," [http://www.econ.lsa.umich.edu/~yanchen/papers/99srl\\_mech.pdf](http://www.econ.lsa.umich.edu/~yanchen/papers/99srl_mech.pdf).
- Dekel, Eddie and Faruk Gul [1997] "Rationality and Knowledge in Game Theory," in *Advances in economics and econometrics: Theory and applications, Seventh World Congress, Volume 1, Econometric Society Monographs, no. 26*, David M. Kreps and Kenneth F. Wallis, (eds.). Cambridge, New York and Melbourne: Cambridge University Press, 87-172.
- Dekel, Eddie, Drew Fudenberg and David K. Levine [1999] "Payoff Information and Self-Confirming Equilibrium," *Journal of Economic Theory*; 89(2), 165-85.
- Fudenberg, Drew and David M. Kreps [1988] "A Theory of Learning, Experimentation, and Equilibrium in Games," unpublished mimeo.
- Fudenberg, Drew and David M. Kreps [1995] "Learning in Extensive-Form Games. I. Self-Confirming Equilibria," *Games and Economic Behavior*; 8(1), 20-55.
- Fudenberg, Drew and David K. Levine [1993] "Self-Confirming Equilibrium," *Econometrica*; 61(3), 523-45.
- Fudenberg, Drew and David K. Levine [1997] "Measuring Player's Losses in Experimental Games," *Quarterly Journal of Economics*, 112, 479-506.



- Harsanyi, John [1967- 1968] "Games with incomplete information played by Bayesian players," *Management Science* 14, 159-182, 320-334, 486-502.
- Jackson, M. and E. Kalai [1997] "Social Learning in Recurring Games," *Games and Economic Behavior* 21, 102-134.
- Jordan, James [1995] "Bayesian Learning in Repeated Games" *Games and Economic Behavior*, 9(1), 8-20.
- Kalai, Ehud and Ehud Lehrer [1993] "Rational Learning Leads to Nash Equilibrium," *Econometrica*; 61(5), 1019-45.
- Mitropoulos, Atanasios [2001] "Learning under minimal information: An experiment on mutual fate control," *Journal of Economic Psychology*, 22(4), 523-557.
- Jörg Oechssler and Burkhard C. Schippers [2002] "Can You Guess the Game You're Playing?" (Forthcoming in GEB),  
<http://www.bgse.uni-bonn.de/%7Eburkhard/cyg5.pdf>.
- Piketty, Thomas [1995] "Social Mobility and Redistributive Politics" *Quarterly Journal of Economics*; 110(3), 551-84.
- Rubinstein, Ariel and Asher Wolinsky [1994] "Rationalizable Conjectural Equilibrium: Between Nash and Rationalizability," *Games and Economic Behavior*; 6(2), 299-311.
- Spector, David [2000] "Rational Debate and One-Dimensional Conflict," *Quarterly Journal of Economics*; 115(1), 181-200.